*Research Article*

# *A simple method to handle non-response bias in school surveys on drug use*

Stéphane Legleye[a], François Beck[a,b], Stanislas Spilka[a,c]

[a] CESP, INSERM U1018, Université Paris-Saclay, Villejuif, France

[b] Direction of prevention and health promotion in Public Health France

[c] Observatoire français des drogues et toxicomanies

Short Title: to be used as running head

Non-response in school surveys

Corresponding Author: Stéphane Legleye, CESP, INSERM U1018, Université Paris-Saclay; Tel: (33 6) 71735829; E-mail: stephane.legleye@inserm.fr

Number of Tables: 6.

Number of Figures: 0.

Word count: 4905.

**Keywords: school survey ; non-response bias ; drug use ; methods**

# *A simple method to handle non-response bias in school surveys on drug use*


## Abstract

**Background:** Studies have proved the positive link between truancy and substances use in school surveys. In spite of this, no adapted weighting treatment is generally provided; even when the share of missing and truant pupils is high, and all drug use estimates are biased downward. The necessary data can be collected: on one side, individual current drug use and past episodes of absence and truancy of the respondents; on the other, the count of the presents and absents the day of the survey, including truants, in each class. However, the nature of these data prevents any classical modelling of the survey response without additional assumptions. **Methods:** We review one method proposed in 2002 by Guttmacher and al. that uses only the individual data and propose two methods that combine both kind of data and in which we can distinguish or not between truancy and legitimate absence. We apply them to the French release of the 2015 Espad survey (European survey project on alcohol and other drugs). The theoretical number of pupils was n=7166; 981 were absent (including 359 truants), while 178 were discarded because of the poor quality of their questionnaires and 6007 were considered final respondents. Assumptions, point estimates and variances are compared. **Results:** Guttmacher' method is not conceptually valid and can lead to irrelevant corrections with high variances. Our estimate of cannabis regular use is 8.6% (std=0.75) instead of 7.7% (std=0.67), that is a non-response bias of circa 14%. C**onclusion:** The proposed approach relies on simple and plausible assumptions; it is preferable to any speculative consideration about the magnitude of the underestimation yielded by the classical weighting procedures. Survey designers should evaluate and discuss the potential bias of their surveys and eventually correct it.



## Keywords: school survey ; non-response bias ; drug use ; methods

## Introduction

Non-response bias is a major concern for survey designers. It depends directly on the difference in outcome between respondents $\bar{Y}_r$ and non-respondents $\bar{Y}_{nr}$, and on the response rate $\bar{R}$ as shown in equation (1):

(Eq. 1): $B = (\bar{Y}_r - \bar{Y}_{nr}) \times (1 - \bar{R})$, where the subscript (r, nr) identifies the respondents and non-respondents, respectively. As $\bar{Y}_{nr}$ is unknown by definition, this bias cannot be estimated without exogenous data or strong hypotheses. A stochastic approach of non-response yields the following equation (1):

(Eq. 2): $B \cong std(R) \times std(Y) \times \sigma(Y, R)/\bar{R}$, where R is the true response propensity *std* is the standard deviation and $\sigma(Y, R)$ the linear correlation coefficient between the two. The key parameters are $\sigma(Y, R)$ and $\bar{R}$. As R is unknown (and has to be estimated with assumptions), the overwhelmingly observed indicator is thus the response rate $\bar{R}$, as in (Eq. 1): the higher $\bar{R}$, the lower the bias.

In general population survey, response rate is considered poorly correlated to the non-response bias (2, 3). But non-response bias can arise when the topic of the survey by itself relates to the response to the survey, when the missing units contribute to a disproportionally high share of the global estimate; this is what reflects $\sigma(Y, R)$ in formula (2). Classical examples include surveys on drug use (in which drug users may be hard to reach). In such situations, classical weighting procedures like post-stratifications and calibrations are unable to correct for non-response bias, that is called non-ignorable (4).

*School surveys on drug use: the importance of school skipping*

School surveys are often considered as immunised to non-response bias because their response rates $\bar{R}$ are usually very high. It is true that almost all present pupils respond to the survey questionnaire, but it is nevertheless not uncommon that roughly 10%-15% of the pupils are missing (5, 6) (page 18). Do the drug use levels of these missing pupils differ so much from those of the respondents that specific statistical treatments are needed to provide unbiased drug use estimates? The response is probably yes for two reasons. First, among respondents, there are strong positive correlations between the number of truancy days (absence without justification) during the past 30 days and the number of times pupils have used alcohol, tobacco and cannabis. This is acknowledged in each of the 35 participating countries (7) (page 198) in the 2003 report of the ESPAD survey (European survey project on alcohol and other drugs). Similarly, positive correlations between the legitimate (with a parental justification) and illegitimate absence (without parental justification) and drug use have also been documented in US school surveys for the period 2000-2008 (8). In fact, the link between school skipping, truancy or disengagement from school and drug use has been documented in western countries (9-13), as well as in other cultural contexts (14).

Second and most important reason, specific interview protocols have led to the conclusion that truants really use more drugs than the others: this was observed in the USA as early as 1975 (15-17) and also in Switzerland in the late 90's (18).

The classical estimates of drug use prevalence in school surveys may thus be underestimated without any appropriate weighting. The problem is old (19) but rarely treated: to our knowledge, none of the classical school surveys on drug addresses this topic directly: Monitoring the future (20), ESPAD (6), HSBC –Health behaviour in school aged-children- (21). To our knowledge, only the study by Guttmacher et al. (16) addresses this issue directly.

Of course, there are other forms of non-response than the absence of a pupil the day of a survey. Non-response can arise at the school level or at the class level; parents can refuse the participation of their children and some questionnaires are also discarded because of their poor quality or incompleteness. These cases represented less than 1% of the pupils in the 2015 ESPAD survey (Guttormsson, Leifman et al. 2016) and cannot bias the estimates because they do not relate to drug use.

*One or two-step weighting procedure*

Classical weighting techniques comprise one-step and two-steps procedures. The most classical is the former for which post-stratifications or calibrations (22) are used. The principle is to reweigh the responding units to obtain the true totals of some variables of the target population, hopefully linked to the variables of interest of the survey. These true totals for these variables are to known by other means: administrative data, census, large and reliable survey samples etc. The second method has two steps: first correcting the probability of being respondent (Total non-response correction: TNRC) then using post-stratification or calibration. It is often more efficient because the first step can use variables, linked to response as well as to the survey outcomes (23), that cannot be considered for post-stratification (24). Reports of absences and truancy by the respondents are good candidates for this purpose.

*The impossibility to model non-response in school surveys*

TNRC is classically achieved through a modelling of the probability of being respondent in order to obtain the total of the target population (by using variables relating to the outcomes and the probability of response). This implies a clear definition of the response to the survey and having the same variables for the respondents and for the non-respondents (23). Generally, the only available data for the absents (non-respondents) in school surveys is defined at the class level: numbers of missing pupils by motive (truancy or not) collected by the survey supervisors, whereas no record of absences and truancy of the missing pupils is available. The available data for the respondents is their reports of missed school days (usually in the last 30 days) by motive (truancy or not). The incommensurability of the two kinds of data prevent their use in any classical modelling. Unfortunately, this prevents also the use of the Heckman method that is specifically designed for the none-ignorable non-response bias (25, 26).

This situation is thus paradoxical, considering that all the relevant variables seem available to reweigh the respondents. The number of absences of the missing pupils is the only information that lacks.

*Attempts to correct for the non-response bias*

At first view, we do not need the class-level information: the probability of a unit to be absent depends only on its characteristics. The reported history of absences of the respondents can be used to approximate their probability to respond to the survey. If a respondent pupil reports 5 days of absence within the last 30 days, one can estimate that he/she was present only 15 days in the average 20 school days during this period. As a consequence, its weight has to be multiplied by a factor 20/15=1.33. This is the procedure followed by (27) and by Guttmacher et al. (2002). This approach relies on a reasonable implicit assumption: the day of the data collection can be considered at random and independent of the pattern of absences of the pupils (see *Limitations*). The idea dates back to 1949 (28) where it was proposed as a mean to avoid multiple visits in face-to-face surveys.

In the case of a school survey, this approach has three problems, however. First, it uses the records of past absences of the respondents without considering the numbers of respondent pupils the day of the data collection, that is the amount of lost information due to missing pupils: without any link between the two, this is not a genuine modelling. Second, because of this, it can lead to irrelevant increases of the weighted number of respondents. Imagine for example a class of 30 pupils where none is absent and all respond to the survey but where 10 pupils report each 10 days of past absence: the weighted number of pupils in the class will be $T=20 + 10\times20/10=40$. The only option (that is not used by Guttmacher et al.) is to reweigh the pupils, that will lead the 20 presents without any past absence to get a weigh below 1 without any justification.

*Objectives*

We propose two simple weightings used as a total non-response correction. Our approach accounts for the probability of being present in class a given day, and for the numbers of present and absent pupils the day of the survey data collection: it assumes a direct link between the two kinds of information. The first version does not distinguish between truancy and legitimate absences while the second one does, in accordance to the literature that emphasizes the prominent role of truancy on drug use. Both methods are compared to Guttmacher's technique and applied to real survey data. Underlying assumptions are discussed.

## Materials and Methods

*The 2015 French ESPAD survey*

ESPAD is coordinated by the Swedish Council for Information on Alcohol and Other Drugs (CAN). It is a 4-years school survey targeting the 15-16 years old of European countries since 1995 (www.espad.org). The main goal of this project is to monitor drug use and to allow comparisons between countries, using a standardised questionnaire and methodology. The French ESPAD survey is conducted by the French monitoring centre for drugs and drug addiction since 1999. The data collection took place in March 2015.

*Sampling*

The sampling frame is the national list of secondary schools, classes and pupils (including their gender and age) from the Ministry of Education drawn in January. The sampling design consists on a stratified sampling of classes; strata are combinations of academic field (8 categories), educational sector (private or public) and city size. In each selected school of each stratum, two classes were selected, with unequal probabilities.

*Survey protocol*

An advance letter informs the parents that a survey on health and lifestyle will take place within a few weeks: the precise topic and the day of the data collection are not given. The teachers have to keep the secret and not to communicate the topic and the day to the pupils. The data collection consists in a pen and paper survey during one teaching class (50 mn). All questionnaires, filled in or not, are placed in an envelope at the end of the school class that is sealed and directly sent to the data capture centre. The data collection is supervised by a professional supervisor whose role was to introducing the survey, ensuring anonymity and confidentiality as well as showing some example of the future use of the data by the researchers. He/she also fills in a classroom report. Questionnaires with high item non-response rate (>50%) or of poor quality are discarded.

*Absence and truancy*

On the first page of the ESPAD questionnaire, respondents are asked to report the number of days they missed school in the last 30 days according to three motives: illness, truancy (no motive) and other reasons. The response scale is 0 day, 1, 2, 3-4, 5-6, 7+ days from which we derived the total number of missed days by motive: 0, 1, 2, 3.5, 5.5 and 7.5. The three motives were aggregated in two categories: illegitimate absences (i.e. truancy) and legitimate absences (illness and other reasons).

The classroom report contains the number of present and absent pupils by gender. Three categories of absences are distinguished: legitimate (there is a parental proof), uncertain (there is a claim by the classmates that the absence is legitimate), illegitimate/truancy (no parental proof and no claim). We distinguished simply the illegitimate absences (truancy) from the others.

*Definition of respondents and non-respondents*

We consider as respondents the present pupils whose questionnaire was retained as valid: missing pupils and respondents with discarded questionnaires were considered as non-respondents, the latter as missing with a legitimate motive.

*Outcomes*

The variables of interest are dichotomous indicators: alcohol and cannabis regular use (at least 10 uses in the last 30 days) as well as tobacco daily smoking (at least one cigarette a day in the last 30 days). These binary indicators are key variables in the French monitoring of drug use in youth.

*Missing values*

Missing values in the report of the past absences/truancy of the respondents were imputed with a random hotdeck procedure considering class, gender, being discarded or not.

*Total non-response corrections*

We implement three TNRC methods that aim to correct the sampling weight in order that the weighted respondent sample have the same size than the theoretical number of pupils.

1/ the Guttmacher method (see above). The formula is:

$$p1_i = W_i \times \frac{20}{20 - N(missed\ school\ days)_i}.$$

*N(missed school days)*$_i$ is the number of missed school days (for any reason) in the last 30 days reported by the respondent $i$ (see *Methods*) and $W_i$ is the sampling weight. We truncated *N(missed school days)*$_i$ to 19 (it could go up to 22.5) as in Guttmacher (2002) because the respondents were by definition present the day of the survey. Because the weights were sometimes very high, we truncated them at the 99$^{th}$ percentile.

2/ In method 2, we consider the number of past absences reported by the respondents as well as the number of non-respondents. The respondents with reported absences will represent themselves plus a proportion of the missing pupils, defined by the ratio of their number of past absences and the total number of days of absences among the respondents. The formula is:

$$p2L_i = W_i + W_i \times C_0(any\ absence)_i$$

$$p2L_i = W_i +$$

$$W_i \times boy \times T(absent\ boys\ )_L \times \frac{N(missed\ school\ days)_i}{T(missed\ school\ days\ reported\ by\ boys)_L} +$$

$$W_i \times girl \times T(absent\ girls\ )_L \times \frac{N(missed\ school\ days)_i}{T(missed\ school\ days\ reported\ by\ girls)_L}$$

where $T(absent\ boys/girls)_L$, is the number of boys/girls that are non-respondents the day of the data collection, $N(missed\ school\ days)_i$ has been defined above and $T(past\ missed\ school\ days)_L$ is the total number of reported missed school days among respondents in the stratum L. If there is no absent the day of the data collection, then $C_0$=0 and $p2L_i$=$W_i$. If no respondent reports a past absence, then p2$L_i$=1 (this did not happen in our case).

3/ In method 3, we extend method 2 by distinguishing the type of absences reported by the respondents and recorded by the supervisors during the survey:

$p3L_i = W_i + W_i \times C_1(legitimate\ absence)_i + W_i \times C_2(truancy)_i$, with:

$$C_1(legitimate\ absence)_i =$$

$$boy \times T(absent\ but\ not\ truant\ boys\ )_L \times \frac{N(missed\ but\ not\ skipped\ days\ )_i}{T(missed\ but\ not\ skipped\ days\ of\ boys\ )_L}$$

$$+ girl \times T(absent\ but\ not\ truant\ girls)_L \times \frac{N(missed\ but\ not\ skipped\ days\ )_i}{T(missed\ but\ not\ skipped\ days\ of\ girls)_L}$$

$$C_2(truancy)_i =$$

$$boy \times T(truant\ boys)_L \times \frac{N(skipped\ days)_i}{T(skipped\ days\ of\ boys)_L}$$

$$+girl \times T(truant\ girls)_L \times \frac{N(skipped\ days)_i}{T(skipped\ days\ of\ girls)_L}$$

$N(skipped\ days)_i$ and $N(missed\ but\ not\ skipped\ days)_i$ are the numbers of skipped days (truancy) and the number of past missed school days for a legitimate motive reported by pupil *i*, respectively, while $T(skipped\ days\ of\ boys)_L$ and $T(missed\ but\ not\ skipped\ days\ of\ boys)_L$ are the corresponding totals in stratum L. If there is no truant the day of the data collection within the considered stratum L, then $C_2$=0; similarly, if there is no absent (legitimate), then $C_1$=0.

Methods 2 and 3 consider the observed correlation between the reported levels of drug use and the probability of response (estimated by the number of missed school days as in Guttmacher' method) in a manner that is analogous to equation (Eq. 2). The truncation of the reported number of missed school days is unnecessary (whereas it is needed in the Guttmacher' method) because the weights are normalised by construction; weights were not truncated for the same reason. Methods 2 and 3 allow considering gender as a determinant of truancy and drug use.

*Underlying assumptions*

In all methods, we assume that there is no self-selection relating to the survey: the day of the survey is at random and no pupil chose to be absent because of the survey (H1). Second, like in every weighting technique, we assume that the unobserved drug uses of the absents or truants can be estimated by the observed drug uses of the respondents (H2), conditionally to specific reports of past absences or truancy episodes (hypothesis of conditional exchangeability). In the Guttmacher' original technique, exchangeability is assumed to hold without condition, whatever the specificities of the pupils: sex, type of school, educational sector etc. On the opposite, using the number of missing pupils at the sampling stratum is more interesting because strata encompass the educational specificities shared by a lot of classes, that relate to the patterns of absence from school and of drug use (see Table 3). We also assume that a missing school day reported by a respondent would be have been (counterfactually) recorded as an absence the by the survey supervisor (H3). H3 makes the modelling of the total non-response possible.

In method 3, we add two additional assumptions. H4: the illegitimate (respectively legitimate) absences, either reported by the respondents or recorded by the survey supervisor belong to the same category. That is, we assume that any reported truancy episode has been recorded as such by a professor during a regular class and would counterfactually be recorded as such by the survey supervisor. H4 is a natural extension of H3. H5: we precise the conditional exchangeability assumption (H2). We assume that the respondents who report truancy episodes can represent the missing pupils without parental justification the day of data collection (i.e. the current truants) regarding drug use in a better way than those who report legitimate absences only.

*Statistics*

All statistics (weighted or not) were computed using the sampling design (strata and class as cluster) to get unbiased estimates of the standard deviation using the PROC SURVEYFREQ in SAS V9.4.

## Results

*Description of the sample*

The selected sample of the 2015 French ESPAD survey comprised 284 classes and 7,166 pupils among which only 6,185 (86.3%) were present and 981 (13.7%) were absent during data collection (including 16 parental refusals). Among the 981 absent pupils, 359 (36.6%) were truant and 116 (11.7%) had an uncertain status, the other having a parental justification. Only 6,185 questionnaires were filled-in (85.9%), while 6,007 were retained in the final respondent sample (83.8%) because 178 questionnaires had to be discarded due to major incompletion or poor data quality. As a consequence, 1159 pupils (almost 16% of the initial sample) were considered missing (among which 31% were current truants). We considered the 116 absent pupils with an uncertain status as absent with a legitimate motive.

The partial non-response rates in the report of past absences were low in the retained questionnaires (5.4% for illness, 8.4% for truancy and 7.1% for other reasons) and higher in the discarded questionnaires (16%, 17% and 18%, respectively). The mean number of reported missed school days was also higher in the discarded questionnaires than in the retained questionnaires (6.2 vs 2.6 days), as well as the proportion of missed days because of truancy (37% vs 25%). Among the retained questionnaires, boys reported less past missed school days without truancy than girls (average 2.5 vs 2.8, $p<0.001$) but the same average of truancy days: 0.7 vs 0.6, $p=0.5$.

*Reported past absences and drug use*

The Pearson correlation coefficients between the outcomes and the reported absences and truancy of the respondents are shown in Table 1. The coefficients with legitimate absences were lower than with those with truancy. Correlations were weak for regular alcohol use (rho circa 0.03) but stronger for tobacco daily smoking and regular cannabis use (rho close to 0.2). The coefficients for smoking and cannabis with the truancy were somewhat higher among girls than boys. As a consequence, drug use levels were higher in respondents who reported a past absence in the last 30 days than among the others and especially high among those who reported episodes of truancy (Table 2). For example, in girls, tobacco daily smoking prevalence was 15.7% among respondents with no reported absence but 26.9% among the others (19.1% among those with legitimate absences only and 41.0% among those with reported episodes of truancy). Respondents with only legitimate absences and those with no absence at all had very similar drug use levels (Table 2).

Truancy is thus clearly the key parameter for any TNRC procedure. However, the numbers of days of absences (legitimate or not) and of truancy episodes were strongly correlated ($r=0.68$ in boys and $r=0.65$ in girls), justifying trying a TNRC procedure that does not distinguish them.

*How much bias can there be?*

According to equation (Eq. 1), the maximum (but unrealistic) bias would be observed if all non-respondents were drug users. With 16.2% of non-respondents and a (unweighted) prevalence of regular cannabis use among respondents of 7.6%, the true value would be 22.5% (Table 2). Using individual reports of absence allows computing a more plausible estimate. If the unobserved proportion of regular cannabis users among missing pupils was equal to the proportion observed among respondents who reported a past episode of truancy (15.6%), the true proportion of regular users would be 8.9%, 1.3 point (or 17%) above the unweighted value, which may be considered worth correcting the data. However, the true bias remains unknown.

*Strata and drug use*

The correlations between the reported number of past missing school days (any absence) and past skipped days (truancy) and the outcomes varies greatly by stratum (Table 3), as well as the levels of drug use: this is an evidence of the relevance of computing the TNRC of methods 2 and 3 at this level instead as at the global level as in method 1.

*Effects of the TNRC weighting procedures*

The three TNRC methods perform differently in reconstituting the theoretical number of pupils: the Guttmacher' method led to a large overestimation (n=7998.3 vs 7166), even with the truncation of the weights (n=7482.3) whereas methods 2 and 3 yielded the exact total (Table 4: Sum without sampling weight). Results were similar when the sampling weight was taken into account (Table 4; Sum with sampling weight).

The TNRC increased the variance of the weights (measured by the coefficient of variation CV). For method 1 (Guttmacher), the truncation (CV=82.4) yields a much lower variance than the original method (CV=118.9). For methods 2 and 3, the variance was lower because each individual correction contributes only to a share of the missing pupils in the stratum. As expected also, the differences between method 2 and 3 were very small because of the high correlation between the numbers of past absences and of past truancy episodes (r=0.68 in boys and r=0.65 in girls). The final calibration reduced the differences between the methods: the CV varies between 81 (for methods 2 and 3) and 89.7 or 118.9 (for the method 1 with or without truncation). Note that the sampling design contributes a lot to the variance of the weights as the CV for the sampling weight is already 68.7 before calibration and 76.1 after.

Table 5 shows the estimates of outcomes, with the different weighting schemes before calibration. As expected, given the correlations observed in Table 1, levels of tobacco and cannabis uses were more corrected upward than the level of alcohol regular use. As expected again, method 3 yielded only slightly greater estimates than method 2. Method 1 (Guttmacher) yielded the most important corrections, especially when the weights were not truncated (i.e. in the original method): the corrected prevalence of

cannabis regular use was9.5% (std=0.86), that is an increase of 25% compared to the unweighted prevalence (7.6%). According to Table 2 and formula (Eq. 1), it would mean that the proportion of cannabis regular use is 19.5% among the missing pupils: a higher value than the prevalence observed among the respondents who reported episodes of truancy (15.6%). The results obtained with the truncation were more realistic. And so were the results obtained with methods 2 and 3: the corrected prevalence of regular cannabis use is 8.7%, that is a relative increase of 14% compared to the raw estimate.

The results obtained after the final calibration are very similar (Table 6). Before and after calibration, the standard deviations obtained through the Guttmacher's method were higher than those obtained with methods 2 and 3, as suggested by the higher CV of the weights (Table 4). None of the corrected estimates fell outside the confidence interval of the classical estimates.

## Discussion

*Summary of the findings*

To our knowledge, this is the first study comparing different methods aiming to correct the potential non-response bias relating to missing pupils in a school survey. Our approach relies on few simple assumptions and provides estimates of the true values based on all the available information, that is preferable to any speculative consideration about the magnitude of the underestimation yielded by the classical weighting procedures. In addition, the increase of variance is small. Ignoring the amount of lost information due to non-response (described by the number of missing pupils the day of the data collection), the Guttmacher' method (2002) is not a sound modelling and leads to irrelevant corrections with higher variances.

*Limitations*

Our results are based on some strong assumptions that may be challenged.

H1 is common to the three methods. It is reasonable because the topic of the survey and the precise day of the survey are not known in advance by the pupils and their parents: it is difficult to imagine that the absence the day of the survey is caused by the survey itself.

H2 is common to every weighting technique: the respondents can represent the non-respondents given some characteristics relating to their probability of response and the level of the outcomes. In our case, the literature emphasises the role of truancy as a key parameter of drug use. As the truants report higher levels of drug use than the others (15-18), and as the number of truancy episodes were correlated to higher levels of drug use among the respondents of our survey, the assumption seems reasonable. The validity of H2 also requires that the dropouts either present the same drug use level than the other truants or represent only a negligible share of them, which is granted because the share of dropout is very low

at 15-16 years old in France where school is still mandatory (less than 1%). That the respondents with report of past truancy episodes are more similar to truants the day of the data collection (H5) than those without such reports relies on the same basis.

On the opposite, we hypothesise that the pupils with no reported absence or truancy cannot represent effective absent pupils, although they show very similar drug use levels. This increases the weights by construction but should not add bias.

The accuracy of the TNRC procedures relies on the accuracy of the data; this is a prerequisite for our five assumptions but it is especially the case for H3 that is at the core of our approach: its validity relies on the honesty of the respondents. H4 and H5 have the same strength and weakness. If the pupils make up truancy episodes as legitimate absences, the total number of missed school days will be more reliable than the distinct counts of legitimate and illegitimate absences, which is an argument in favour of method 2. Such a trick would be in accordance with a social desirability bias (29). At the class level, it is likely that a proportion of the apparent legitimate absences are in fact illegitimate and reciprocally but this proportion should be low.

The discarded questionnaires have been considered as questionnaires of absent pupils (and not truant pupils), despite showing much higher rates of absence and truancy than the retained questionnaires. This is arbitrary, as they present higher shares of reported truancy; but one reason for this choice is that they were less trustworthy since they showed higher rates of missing values for past absences.

In our case, the correlations between truancy and drug use do not differ much by gender and ignoring gender in methods 2 and 3 would provide very similar results. However, it is important to show that this important determinant of drug use and school attendance can be considered easily.

## Conclusion

Combining the number of missing pupils in class and the individual reports of past missed school days by the respondents allows estimating and correcting the non-response bias in a simple that should be applied in every school survey.

## Declarations

## Ethics approval and consent to participate

Ethics approval was not required for this school survey. The data are anonymous and confidential and protected by the National Committee on Informatics and liberty. The survey was not mandatory; parents could refuse the participation of their children and the pupils could refuse to participate.

## Consent for publication

This manuscript has been seen and approved by all authors, which have been personally and actively involved in substantive work leading to this article, and will hold themselves jointly and individually responsible for its content.

## Competing interests

The authors have no conflicts of interest to declare.

## Funding

The ESPAD survey is funded by the French monitoring centre for drug and drug addiction.

## Authors' contributions

SS designed the survey and was responsible of it. SL designed the study, wrote the first version of the manuscript and computed the statistics. FB revised the manuscript.

## Availability of data and materials

Data and SAS programs are available upon request to the corresponding author.

## References


1. Bethlehem J. Working with Response Probabilities. Journal of Official Statistics. 2020;36(3):647-74.
2. Davern MD, McAlpine T, Beebe J, Ziegenfuss JY, Rockwood TH, Call KT. Are lower response rates hazardous to your health survey? An analysis of three state health surveys. Health services research. 2010;45(5):1324-44.
3. Groves RM. Nonresponse Rates And Nonresponse Bias In Household Surveys. Public Opinion Quarterly. 2006;70(5):646-75.
4. Rubin DB, Little RJA. Statistical Analysis with Missing Data: John Willey & Sons; 2002.
5. Hibell B, Guttormsson U, Ahlström S, Balakireva O, Bjarnason T, Kokkevi A, et al. The 2011 ESPAD Report: Substance Use Among Students in 36 European Countries. Stockholm: CAN; 2012.
6. Guttormsson U, Leifman H, Kraus L, Arpa S, Molinaro S, Monshouwer K, et al. Espad 2015 methodology. Lisbon: EMCDDA; 2016.
7. Hibell B, Andersson B, Bjarnason T, Ahlström S, Balakireva O, Kokkevi A, et al. The ESPAD Report 2003. Alcohol and other drug use among students in 35 European countries. Rapport. Stockholm: CAN (Council for Information on Alcohol and other Drugs); 2004.
8. Gfroerer J, Bose J, Kroutil L, Lopez M, Kan L. Methodological Considerations in Estimating Adolescent Substance Use. Joint Statisticl Meeting; San Siego2012.
9. Wichstrom L. Alcohol intoxication and school dropout. Drug Alcohol Rev. 1998;17(4):413-21.
10. Henry KH, Thornberry TP. Truancy and Escalation of Substance Use During Adolescence. Journal of Studies on Alcohol and Drugs. 2010;71:115-24.
11. Mounteney J, Skutle A. Truancy, alcohol use and alcohol-related problems in secondary school pupils in Norway. Health Education Research. 2010;25(6):945-54.
12. Holtes M, Bannink R, Joosten-van Zwanenburg E, Van As E, Raat H, Broeren S. Associations of truancy, perceived school performance, and mental health with alcohol consumption among adolescents. Journal of School Health. 2015;85(12).
13. Henry KL, Huizinga DH. Truancy's effect on the onset of drug use among urban adolescents placed at risk. Journal of Adolescent Health. 2007;40(4):e9–e17.
14. Yoep N, Kuang Kuay L, Safiza N. Prevalence of Truancy and Its Associated Factors among School-Going Malaysian Adolescents: Data from Global School-Based Health Survey 2012. Psychology. 2016;7:1053-60.
15. Kandel D. Reaching the hard-to-reach: illicit drug use among high school absentees. Addictive diseases. 1975;1(4):465-80.
16. Guttmacher S, Weitzman BC, Kapadia F, Weinberg SL. Classroom-Based Surveys of Adolescent Risk-Taking Behaviors: Reducing the Bias of Absenteeism. American Journal of Public Health. 2002;92:235-7.
17. Swaim RC, Beauvais FE, Chavez L, Oetting ER. The effect of school dropout rates on estimates of adolescentsubstance use among three racial/ethnic groups. American Journal of Public Health. 1997;87(1):51-5.
18. Michaud PA, Delbos-Piot I, Narring F. Silent dropouts in health surveys: are nonrespondent absent teenagers different from those who participate in school-based health surveys? J Adolesc Health. 1998;22(4):326-33.
19. Johnston L, O'Malley PM. Issues of validity and population coverage in student surveys of drug use. . In: Rouse B, Kozel N, Richards L, editors. Self-Report Methods of Estimating Drug Use: Meeting Current Challenges to Validity 57. Rockville, Md: NIDA Research Monograph 1985.
20. Miech RA, Johnston LD, O'Malley PM, Bachman JG, Schulenberg JE, Patrick M. Monitoring the Future national survey results on drug use, 1975–2019. Ann Arbor Michigan: Institute for Social Research The university of Michigan; 2020.

21. Moor I, Winter K, Bilz L, Bucksch J, Finne E, John N, et al. The 2017/18 Health Behaviour in School-aged Children (HBSC) study – Methodology of the World Health Organization's child and adolescent health study. Journal of Health Monitoring. 2020;5(3).
22. Deville J, Sarndal C-E. Calibration estimators in survey sampling. Journal of the American Statistical Association. 1992;87(418):376-82.
23. Little RJA, Vartivarian S. Does weighting for nonresponse increase the variance of survey means? Survey methodology. 2005;31(2):161–68.
24. Haziza D, Lesage E. A discussion of weighting procedures for unit nonresponse. Journal of Official Statistics. 2016;32(1):129-45.
25. Heckman JJ. Sample selection bias as a specication error. Econometrica: Journal of the econometric society. 1979;47(1):153-61.
26. Gallimard J-E, Chevret S, Curis E, Resche-Rigon M. Heckman imputation models for binary or continuous MNAR outcomes and MAR predictors. BMC medical research methodology. 2018;18(90).
27. Bachman JG, O'Malley PM, Johnston LD, Schulenberg JE, Freedman-Doan P. THE MONITORING THE FUTURE EIGHTH GRADE PANEL SURVEY DATA: SAMPLE DESIGN, ADJUSTMENTS FOR PANEL ATTRITION BIASES, AND ASSESSMENT OF MEASUREMENT BIAS. occasional papers. 2001(55).
28. Politz A, Simmons W. An attempt to get the "not at homes" into the sample without callbacks. Journal of the American Statistical Association. 1949;44(245):9-16.
29. De Leeuw ED. Choosing the method of data collection. In: De Leeuw ED, Hox JJ, Dillman DA, editors. International handbook of survey methodology. New York: Lawrence Earlbaum Associates; 2008. p. 117-35.